

# ОЦЕНКА ИНФОРМАТИВНОСТИ ДИАГНОСТИЧЕСКИХ ПАРАМЕТРОВ

Абдуллаев М.У., Бобоёров Р.О.

Ташкентский химико-технологический институт

В целях выбора того или иного диагностического параметра предварительно осуществляют оценку информативности каждого из них. Эта оценка основывается на основных положениях теории информации.

В качестве опыта  $\alpha$  рассмотрим определение состояния системы с возможными несовместными исходами  $S_0, S_1, S_2, \dots, S_n$ , составляющими полную группу, и изменим обозначение  $\alpha$  на  $S$ . В качестве опыта  $\beta$  – выявление значения логической переменной  $Z$  (индекс пока опустим) для одного из диагностических параметров; символ  $\beta$  изменим на  $Z$ . В качестве меры информативности параметра возьмем количество информации, содержащееся в Зотносительно  $S$ :

$$I(S/Z) = H(S) - H(S/Z)$$

Вычисления по этой формуле вполне можно выполнить, имея в качестве исходных данных вероятности  $P(S_0), P(S_1), P(S_2), \dots, P(S_n)$  состояний  $S_0, S_1, S_2, \dots, S_n$ . Они могут быть известны как опытные данные или рассчитаны методами теории надежности. Заметим, что

$$P(S_0) + P(S_1) + \dots + P(S_n) = 1.$$

Кроме того, при вычислениях потребуются условные вероятности состояний  $S_0, S_1, S_2, \dots, S_n$  при наличии событий « $Z = 0$ » и « $Z = 1$ ». Соответствующие расчеты тоже можно выполнить, пользуясь известными формулами теории вероятностей.

Однако значительно проще воспользоваться равенством (П ), согласно которому  $I(S/Z) = I(Z/S)$  и провести вычисления по формуле

$$I(S/Z) = I(Z/S) = H(Z) - H(Z/S).$$

Упрощения обусловлены тем, что согласно структуре диагностической таблицы значения  $Z$  могут быть равны только 0 или 1 и при этом они однозначно определяются состояниями.

Найдем  $H(Z)$  по формуле

$$H(Z) = -(P(Z=0)\log_2 P(Z=0) + P(Z=1)\log_2 P(Z=1)).$$

Входящие в формулу вероятности легко определяются по диагностической таблице на основании известных вероятностей  $P(S_0), P(S_1), P(S_2), \dots, P(S_n)$ . Действительно, вероятность события « $Z=0$ » равна сумме вероятностей состояний, для которых признак  $Z$  обращается в 0. Это те состояния, против которых в строке рассматриваемого параметра стоит символ 0. Обозначим сумму их вероятностей  $\Sigma_0$ . Вероятность события « $Z=1$ » есть сумма вероятностей состояний, для которых в строке параметра стоит символ 1. Обозначим эту сумму  $\Sigma_1$ . Таким образом,

$$H(Z) = -(\Sigma_0 \log_2 \Sigma_0 + \Sigma_1 \log_2 \Sigma_1).$$

Второе слагаемое в формуле обращается в 0, поскольку значение  $Z$  однозначно определяется, если известно состояние системы . Окончательно имеем в качестве меры информативности выбранного параметра

$$I(S/Z) = -(\Sigma_0 \log_2 \Sigma_0 + \Sigma_1 \log_2 \Sigma_1).$$

Эта формула может быть применена к каждому из параметров  $Z_1, Z_2, \dots, Z_m$ , после чего оказывается возможным их сравнение по информативности.

Если для какого-то параметра оказывается, что в его строке стоят только нули, то  $\Sigma 0 = 1$ ; если только 1, то  $\Sigma 1 = 1$ . В обоих случаях для этого параметра  $I(S/Z) = 0$ , он не несет информации о состояниях и должен быть исключен.

Если необходим только выбор наиболее информативного параметра, а оценка количества информации не нужна, то можно пользоваться простым правилом: следует выбирать тот параметр, для которого  $\Sigma 0$  или  $\Sigma 1$  наименее отличаются от значения 0,5. Имеет место равенство

$$\Sigma 0 + \Sigma 1 = 1,$$

Для этого достаточно заменить обозначение  $\Sigma 0$  на  $p$ , а  $\Sigma 1$  – на  $(1-p)$  или наоборот.

Пусть  $P(S0) = 0,8$ ;  $P(S1) = P(S2) = P(S3) = 0,04$ ;

$P(S4) = 0,08$ . Тогда для параметра  $Z1$ :

$$\Sigma 0 = P(S1) + P(S2) + P(S4) = 0,16; \Sigma 1 = P(S0) + P(S3) = 0,84.$$

Для параметра  $Z2$ :

$$\Sigma 0 = P(S1) + P(S2) + P(S3) + P(S4) = 0,2; \Sigma 1 = P(S0) = 0,8.$$

Для параметра  $Z3$ :

$$\Sigma 0 = P(S4) = 0,08; \Sigma 1 = P(S0) + P(S1) + P(S2) + P(S3) = 0,92.$$

И далее

$$I(S/Z1) = 0,6343; I(S/Z2) = 0,7219; I(S/Z3) = 0,4022 \text{ (все величины – в битах).}$$

Наиболее информативным оказался параметр  $Z2$ , исходя из этого следует, что для оценки информативности диагностических параметров целесообразно принять параметр  $Z2$ .

#### Литература

1. Автоматизация типовых технологических процессов и установок: Учеб.для вузов/А.М.Корытин,Н.К.Петров, С.Н.Радимов, Н.К.Шапарев.-2-е изд.,перераб. и доп.- М.:Энергоатомиздат, 1988.-432с.: ил.
2. ГОСТ 27.002-89. Надежность в технике. Термины и определения.-М.: Изд-во стандартов,1989.-37с.
3. ГОСТ 20911-89. Техническая диагностика. Термины и определения.- М.: Изд-во стандартов,1989.-13с.